(2)

SECURITY CLASSIFICATION OF THIS PAGE

AD-A197 323

DTIC ELECTE AUG 1 1 1988 D H

REPORT DOCUMENTATION PAGE

1b. RESTRICTIVE MARKINGS

DTIC FILE COPY

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE
NA

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for Public Release; Distribution Unlimited

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
FSU Technical Report No. M-783

5. MONITORING ORGANIZATION REPORT NUMBER(S)
ARO 23699.19-MA

6a. NAME OF PERFORMING ORGANIZATION
Florida State University

6b. OFFICE SYMBOL (If applicable)

7a. NAME OF MONITORING ORGANIZATION
AFOSR/NM

6c. ADDRESS (City, State and ZIP Code)
Department of Statistics
Tallahassee, FL 32306-3033

7b. ADDRESS (City, State and ZIP Code)
Bldg. 410
Bolling AFB, DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL (If applicable)
NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
DAAL03-86-K-0094

8c. ADDRESS (City, State and ZIP Code)
Bldg. 410
Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.

11. TITLE (Include Security Classification) Identifying Nonlinear Covariate effects in Semimartingale Regression Models

12. PERSONAL AUTHOR(S)
Ian W. McKeague and Klaus J. Utikal

13a. TYPE OF REPORT
Technical

13b. TIME COVERED
FROM ________ TO ________

14. DATE OF REPORT (Yr., Mo., Day)
June, 1988

15. PAGE COUNT
21

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Martingale methods, counting processes, diffusion processes Gaussian random fields, censored survival data, test for independence, time-homogeneity, proportional hazards.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Let X_t be a semimartingale which is either continuous or of counting process type and which satisfies the stochastic differential equation $dX_t = Y_t\,\alpha(t, Z_t)\,dt + dM_t$, where Y and Z are predictable covariate processes, M is a martingale and α is an unknown, nonrandom function. We study inference for α by introducing an estimator for $A(t,z) = \int_0^z \int_0^t \alpha(s,x)\,ds\,dx$ and deriving a functional central limit theorem for the estimator. The asymptotic distribution turns out to be given by a Gaussian random field that admits a representation as a stochastic integral with respect to a multiparameter Wiener process. This result is used to develop a test for independence of X from the covariate Z, a test for time-homogeneity of α, and a goodness-of-fit test for the proportional hazards model $\alpha(t,z) = \alpha_1(t)\,\alpha_2(z)$ used in survival analysis.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Frank Proschan

22b. TELEPHONE NUMBER (Include Area Code)
(904) 644-3218

22c. OFFICE SYMBOL
AFOSR/NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

ARO 23699.19-MA

IDENTIFYING NONLINEAR COVARIATE EFFECTS IN SEMIMARTINGALE REGRESSION MODELS

by

Ian W. McKeague[1] and Klaus J. Utikal[2]

FSU Technical Report No. M-783
USARO Technical Report No. D-103
AFOSR Technical Report No. 88-229

June, 1988

Department of Statistics
The Florida State University
Tallahassee, FL 32306-3033

[1] Research supported by the Army Research Office under Grant DAAL03-86-K-0094.
[2] Research partially supported by the Air Force Office of Scientific Research under Contract No. 88-0040.

AMS 1980 subject classifications. 62M99, 62J02, 62G05.

Key words and phrases. Martingale methods, counting processes, diffusion processes, Gaussian random fields, censored survival data, test for independence, time-homogeneity, proportional hazards.

Abstract

Let X_t be a semimartingale which is either continuous or of counting process type and which satisfies the stochastic differential equation $dX_t = Y_t\,\alpha(t, Z_t)\,dt + dM_t$, where Y and Z are predictable covariate processes, M is a martingale and α is an unknown, nonrandom function. We study inference for α by introducing an estimator for $A(t,z) = \int_0^z \int_0^t \alpha(s,x)\,ds\,dx$ and deriving a functional central limit theorem for ~~the estimator~~. The asymptotic distribution turns out to be given by a Gaussian random field that admits a representation as a stochastic integral with respect to a multiparameter Wiener process. This result is used to develop a test for independence of X from the covariate Z, a test for time-homogeneity of α, and a goodness-of-fit test for the proportional hazards model $\alpha(t,z) = \alpha_1(t)\,\alpha_2(z)$ used in survival analysis.

1. Introduction

Consider a nonlinear semimartingale regression model in which a process X is related to a covariate process Z by

$$X_t = X_0 + \int_0^t \lambda_s \, ds + M_t, \tag{1.1}$$

$$\lambda_t = Y_t \, \alpha(t, Z_t), \tag{1.2}$$

where α is an unknown, bounded, deterministic function, M is a martingale and Y is an indicator process, taking the value 1 when X and Z are under observation, zero otherwise. In the case that X is a counting process, λ and α are called the intensity process and conditional hazard function respectively. If the intensity process is of the form $\lambda_t = \alpha(t) \, Z_t$, we have Aalen's (1978) multiplicative intensity model, for which a well developed theory of hazard rate and integrated hazard rate estimation exists (see the survey article of Andersen and Borgan, 1985). For the nonlinear model (1.2) an estimator $\hat{A}(\cdot, z)$ of the time-integrated conditional hazard function $A(t,z) = \int_0^t \alpha(s,z)\, ds$ at a fixed level z of the covariate Z has been studied by Beran (1981) and Dabrowska (1987) in the survival analysis setting, and by McKeague and Utikal (1987) in the general case. This estimator was used to develop methods of inference for the function $\alpha(\cdot, z)$ at fixed z, based on observation of i.i.d. replicates of (X, Y, Z).

In the present paper we study inference for the entire conditional "hazard" function $\alpha(\cdot,\cdot)$. For that purpose we introduce the estimator

$$\hat{\mathcal{A}}(t,z) = \int_0^z \hat{A}(t,x)\, dx$$

of the time and state integrated hazard function

$$\mathcal{A}(t,z) = \int_0^z \int_0^t \alpha(s,x)\, ds\, dx = \int_0^z A(t,x)\, dx.$$

When X is a continuous process or a counting process Theorem 3.1 gives the weak convergence of the appropriately normalized time and state indexed process $\hat{\mathcal{A}}$ to a Gaussian random field. This is proved by using the results of Bickel and Wichura (1971) to establish tightness. Convergence of the finite dimensional distributions is shown using Rebolledo's (1980) martingale central limit theorem.

In Section 4.1 we propose a test for independence of X from the covariate process Z. Here independence from the covariate means that α is only a function of time. A natural estimator for $\mathcal{A}$ under the hypothesis of independence is given by $\bar{\mathcal{A}}(t,z) = z\, \bar{A}(t)$, where $\bar{A}$ is the Nelson-Aalen estimator. We derive the asymptotic distribution of $\hat{\mathcal{A}} - \bar{\mathcal{A}}$ in Theorem 4.1 and show that a maximal deviation statistic based on $\hat{\mathcal{A}} - \bar{\mathcal{A}}$ yields a consistent test for independence.

In Section 4.2 we propose a test for time-homogeneity, i.e. that $\alpha = \alpha(t,z)$ does not depend on time t. An estimator for $\mathcal{A}$ under the hypothesis of time-homogeneity is given by $\mathcal{A}^*(t,z) = t\, \hat{\mathcal{A}}(1,z)$. A maximal deviation test statistic based on $\hat{\mathcal{A}} - \mathcal{A}^*$ is shown to yield a consistent test for time-homogeneity.

In Section 4.3 we develop a goodness-of-fit test for the "proportional hazards" model $\alpha(t,z) = \alpha_1(t)\, \alpha_2(z)$, where $\alpha_1(t)$ and $\alpha_2(z)$ are arbitrary unknown functions. This model has been studied by Thomas (1983), Tibshirani (1984), Hastie and Tibshirani (1986) and O'Sullivan (1986a, 1986b) in the survival analysis context (where it is a generalization of Cox's (1972) proportional hazards

model). These authors propose various estimators for the log relative risk function $\log\alpha_2$, where α_2 is assumed to be positive, but, except for O'Sullivan (1986a), who finds a rate of convergence for his estimator, they do not provide any asymptotic theory. We introduce $\hat{A}(1,\cdot)$ as an estimator of the cumulative relative risk function $A_2(\cdot) = \int_0^{\cdot} \alpha_2(x)\,dx$ and find its asymptotic distribution. Technical lemmas used in the proofs of our main results are given in Section 5.

2. Preliminaries

Let $(\Omega, \mathcal{F}, P)$ denote a complete probability space, $(\mathcal{F}_t, t \in [0,1])$ a nondecreasing, right-continuous family of sub-σ-fields of $\mathcal{F}$, where $\mathcal{F}_0$ contains all P-null sets in $\mathcal{F}$. All processes are indexed by $t \in [0,1]$. The process $(M_t, \mathcal{F}_t)$ is assumed to be a zero-mean L^4-martingale with sample paths in Skorohod space $D[0,1]$. The quadratic characteristic of M will be denoted by $\langle M \rangle$ and its quadratic variation by $[M]$. The processes Y and Z are assumed to be predictable, with Y an indicator process. For simplicity, Z is supposed to be scalar valued. The processes X, Y, Z and M are related by (1.1) and (1.2) which can be written in the form

$$dX_t = Y_t\,\alpha(t, Z_t)\,dt + dM_t. \tag{2.1}$$

We assume that

$$\langle M \rangle_t = \int_0^t \gamma(t, Z_s, Y_s)\,ds, \tag{2.2}$$

where γ is a bounded, measurable function. Note that if X is a counting process we have

$$\gamma(t,z,y) = \alpha(t,z)\,y. \tag{2.3}$$

Let $W = (W(t,z),\ (t,z) \in [0,1]^2)$ be a two-parameter Wiener process, i.e. a Gaussian process with zero mean and $E\,W(t,z)\,W(t',z') = \min(t,t')\min(z,z')$. Let $\int_0^t \int_0^z \phi(s,x)\,dW(s,x)$ denote a continuous version of the Wiener integral of a function $\phi \in L^2([0,1]^2, ds\,dx)$ defined by Ito (1951), Wong and Zakai (1974) and Bass (1988). The estimators and test statistics that we shall introduce have asymptotic distributions which can be represented in terms of stochastic integrals of this type.

We make the following assumptions:

(A1) For each t, the random vector (Z_t, Y_t) is absolutely continuous with respect to the product of the Lebesgue and counting measure. Denote the corresponding density by $f_{Z(t)Y(t)}(z,y)$.

(A2) $f_{Z(t)Y(t)}(z,1)$ is bounded away from zero.

(A3) $f_{Z(t)Y(t)}(z,1)$ is continuous as a function of t and z.

(B1) $\gamma(t,z,y)$ is a continuous function of t and z for each fixed y.

(B2) α is Lipschitz, i.e. there exists a constant K such that

$$|\alpha(t_1, z_1) - \alpha(t_2, z_2)| \le K\sqrt{(t_1 - t_2)^2 + (z_1 - z_2)^2}$$

for all t_1, t_2, z_1, z_2.

Let $C_2 = C([0,1]^2)$ denote the space of continuous functions on the unit square equipped with the supremum norm $\|\cdot\|$. Let D_2 denote the extension of the space $D[0,1]$ to functions on $[0,1]^2$, as described in Neuhaus (1971).

3. Estimation of $\mathcal{A}$

For simplicity we restrict the region over which $\mathcal{A}$ is to be estimated to $[0,1]^2$. For each $n \geq 1$, let $I_1^{(n)}, \ldots, I_{d_n}^{(n)}$ be the partition of the interval $[0,1]$ defined by

$$I_r^{(n)} = \left[\frac{r-1}{d_n}, \frac{r}{d_n}\right), \qquad r = 1, \ldots, d_n - 1$$
$$I_{d_n}^{(n)} = \left[1 - \frac{1}{d_n}, 1\right],$$

where d_n is an increasing sequence of positive integers. The superscript n will usually be suppressed in the notation, for example we shall write $I_1, \ldots, I_{d_n}$ instead of $I_1^{(n)}, \ldots, I_{d_n}^{(n)}$. Let (X_i, Y_i, Z_i, M_i), $i = 1, \ldots, n$ denote copies of the generic processes defined above, where only M_i is not observable and the corresponding filtrations are independent. Define

$$X_r^{(n)}(t) = \sum_{i=1}^{n} \int_0^t I\{Z_i(s) \in I_r\} Y_i(s)\, dX_i(s), \tag{3.1}$$

$$Y_r^{(n)}(s) = \sum_{i=1}^{n} I\{Z_i(s) \in I_r\} Y_i(s), \tag{3.2}$$

$$\hat{A}(t,z) = \int_0^t \frac{1}{Y_r^{(n)}(s)} dX_r^{(n)}(s), \text{ for } z \in I_r,$$

where $1/0 \equiv 0$. Since $\mathcal{A}(t,z) = \int_0^z A(t,x)\, dx$, we propose to estimate $\mathcal{A}$ by

$$\hat{\mathcal{A}}(t,z) = \int_0^z \hat{A}(t,x)\, dx.$$

The asymptotic distribution of $\hat{\mathcal{A}}$ is given by the following result.

THEOREM 3.1. Suppose that A1-A3, B1, B2 hold, $d_n^2/n \to \infty$, $d_n = o(n^\delta)$ for some $\delta \in (1/2, 1)$ and X is a counting process or has continuous sample paths. Then

$$\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}) \xrightarrow{\mathcal{D}} m$$

in D_2 as $n \to \infty$, where

$$m(t,z) = \int_0^t \int_0^z \sqrt{h(s,x)}\, dW(s,x),$$
$$h(s,x) = \frac{\gamma(s,x,1)}{f_{Z(s)Y(s)}(x,1)}.$$

REMARK. The process m is a continuous Gaussian random field with mean zero and covariance function

$$\mathrm{Cov}(m(t_1,z_1), m(t_2,z_2)) = \int_0^{z_1 \wedge z_2} \int_0^{t_1 \wedge t_2} h(s,x)\, ds\, dx.$$

Proof of Theorem 3.1. Define

$$dM_r^{(n)}(s) = \sum_{i=1}^{n} I\{Z_i(s) \in I_r\} Y_i(s)\, dM_i(s), \tag{3.4}$$

$$\alpha_r^{(n)}(s) = \sum_{i=1}^{n} I\{Z_i(s) \in I_r\} Y_i(s) \alpha(s, Z_i(s)). \tag{3.5}$$

Then by (2.1), (3.1)

$$dX_r^{(n)}(s) = \alpha_r^{(n)}(s)\, ds + dM_r^{(n)}(s). \tag{3.6}$$

Also define the processes

$$\widehat{M}^{(n)}(t,z) = \sqrt{n}\sum_{r=1}^{d_n}\int_0^z\int_0^t \frac{1}{Y_r^{(n)}(s)}\, dM_r^{(n)}(s)\, I(x\in I_r)\, dx, \tag{3.7}$$

$$\hat{\mathcal{A}}_p(t,z) = \sum_{r=1}^{d_n}\int_0^z\int_0^t \frac{\alpha_r^{(n)}(s)}{Y_r^{(n)}(s)}\, ds\, I(x\in I_r)\, dx, \tag{3.8}$$

$$\widetilde{M}^{(n)}(t,z) = \frac{\sqrt{n}}{d_n}\sum_{r=1}^{[zd_n]}\int_0^t \frac{1}{Y_r^{(n)}(s)}\, dM_r^{(n)}(s). \tag{3.9}$$

Here and in the sequel, any summation over $r = 1, \ldots, [zd_n]$ is defined to be zero when $[zd_n] = 0$. Now $\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}) = \sqrt{n}(\hat{\mathcal{A}}_p - \mathcal{A}) + \widehat{M}$. Lemma 1 gives $\sqrt{n}\|\hat{\mathcal{A}}_p - \mathcal{A}\| \xrightarrow{P} 0$. To complete the proof we need to show that $\widehat{M}^{(n)} \xrightarrow{\mathcal{D}} m$ in D_2.

Suppose that $\widetilde{M}^{(n)} \xrightarrow{\mathcal{D}} m$ in D_2. Define a linear map $\pi_n : D_2 \to D_2$ by $\pi_n(f)(t,z) = f(t, z_{r-1}) + d_n(z - z_{r-1})f(t, z_r)$ for $z \in I_r$, where $z_r = r/d_n$. Here $\pi_n(f)(t,\cdot)$ is a piecewise linear approximation to $f(t,\cdot)$ based on the points z_r, $r = 1, \ldots, d_n$, for each t. Note that $\widehat{M}^{(n)} = \pi_n(\widetilde{M}^{(n)})$. Also, appealing to a D_2 version of Lemma 4.1 of McKeague (1988), we have $\pi_n(\widetilde{M}^{(n)}) \xrightarrow{\mathcal{D}} m$ in D_2, where we have used the fact that m has its sample paths in C_2.Thus $\widehat{M}^{(n)} \xrightarrow{\mathcal{D}} m$ in D_2. All that remains to be proved is that $\widetilde{M}^{(n)} \to m$ weakly in D_2. This will be established by showing that $\{\widetilde{M}^{(n)}, n \geq 1\}$ is tight in D_2 and the finite dimensional distributions of $\widetilde{M}^{(n)}$ converge weakly to those of m.

Denote the increment of $\widetilde{M}^{(n)}$ over the rectangle $(s,t] \times (x,y]$ by $\widetilde{M}^{(n)}((s,t] \times (x,y]) = \widetilde{M}^{(n)}(t,y) - \widetilde{M}^{(n)}(s,y) - \widetilde{M}^{(n)}(t,x) + \widetilde{M}^{(n)}(s,x)$. Tightness is established by checking some product moment conditions of Bickel and Wichura (1971) for the increments of $\widetilde{M}^{(n)}$ over certain neighbouring rectangles:

$$E\,(\widetilde{M}^{(n)}((s,t] \times (x,y]))^2(\widetilde{M}^{(n)}((s,t] \times (y,z]))^2 \leq K\,(t-s)^{\frac{3}{2}}(y-x)\,(z-y)$$

and

$$E\,(\widetilde{M}^{(n)}((s,t] \times (x,y]))^2(\widetilde{M}^{(n)}((t,u] \times (x,y]))^2 \leq K\,(t-s)^{\frac{1}{2}}\,(u-t)\,(y-x)^2.$$

This is done in Lemmas 2 and 3.

To show convergence of all finite dimensional distributions it suffices to show that for any $0 \leq z_0 < \cdots < z_p \leq 1$, $p \geq 1$

$$(\widetilde{M}(\cdot, z_j) - \widetilde{M}(\cdot, z_{j-1}))_{j=1}^p \xrightarrow{\mathcal{D}} (m(\cdot, z_j) - m(\cdot, z_{j-1}))_{j=1}^p$$

in $D[0,1]^p$, where $D[0,1]^p$ is the product of p copies of $D[0,1]$. This can be done using a p-variate version of Rebolledo's (1980) martingale central limit theorem, as given by Aalen (1977) and Andersen and Gill (1982, Theorem I.2) in the counting process case. The processes $\widetilde{M}(\cdot, z_j) - \widetilde{M}(\cdot, z_{j-1})$, $j = 1, \ldots, p$ are orthogonal square integrable martingales and by Lemma 4

$$\langle \widetilde{M}(\cdot, z_j) - \widetilde{M}(\cdot, z_{j-1})\rangle_t \xrightarrow{P} \langle m(\cdot, z_j) - m(\cdot, z_{j-1})\rangle_t,$$

for each t, $j = 1, \ldots, p$. That completes the proof for the continuous case. In the counting process case we also need to check the Lindeberg condition (cf. Andersen and Gill's (I.4) with $l = r$, $i = j$, $n = d_n$)

$$\sum_{r=1}^{d_n} \int_0^1 H_{jr}^{(n)}(s)^2 I\{|H_{jr}^{(n)}(s)| > \epsilon\} d\langle M_r^{(n)}\rangle_s \xrightarrow{P} 0, \tag{3.10}$$

for all $\epsilon > 0$, where

$$H_{jr}^{(n)}(s) = \begin{cases} \frac{\sqrt{n}}{d_n} \frac{1}{Y_r^{(n)}(s)} & \text{if } [z_{j-1} d_n] < r \le [z_j d_n] \\ 0 & \text{otherwise,} \end{cases}$$

This is done in Lemma 6.

Confidence sets for $\mathcal{A}$

In order to apply Theorem 3.1 to obtain Kolmogorov-Smirnov type confidence sets for $\mathcal{A}$ of the form $\{\mathcal{A}: \sqrt{n} \sup_{t,z} |\hat{\mathcal{A}}(t,z) - \mathcal{A}(t,z)| \le c\}$ we would need to determine the quantiles of $\tau = \sup_{t,z} |m(t,z)|$. In the time-homogeneous case, considered below, it is possible to use existing tables. In the general case, the representation of m in terms of the Brownian sheet process W gives a way to obtain such quantiles by simulation. We shall only consider this in the counting process case, but the continuous case is similar. First estimate the function $H(t,z) = \int_0^t \int_0^z h(s,x)\, dx\, ds$ by

$$\hat{H}(t,z) = \frac{n}{d_n^2} \sum_{r=1}^{[zd_n]} \int_0^t \frac{1}{(Y_r^{(n)}(s))^2} dX_r^{(n)}(s)$$

and then estimate h by

$$\hat{h}(t,z) = \frac{1}{b_n^2} \int_0^1 \int_0^1 K\Big(\frac{t-s}{b_n}\Big) K\Big(\frac{z-x}{b_n}\Big)\, d\hat{H}(s,x),$$

where K is a bounded, nonnegative kernel function with compact support, integral 1 and b_n is a bandwidth parameter, $b_n \to 0$. The following result, which is proved in Lemma 10(a), shows that $\hat{h}$ is an L^2-consistent estimator of h.

PROPOSITION 3.2. Suppose that X is a counting process, the assumptions of Theorem 3.1 hold, $d_n b_n^2 \to \infty$ and K is Lipschitz. Then $E \int_0^1 \int_0^1 |\hat{h}(t,z) - h(t,z)|^2 dt\, dz \to 0$.

The process m, with $\hat{h}$ in place of h, could then be simulated to obtain approximate quantiles for τ. Using Proposition 3.2 it can be shown (cf. the proof of Proposition 4.2) that this procedure leads to asymptotically correct confidence sets for $\mathcal{A}$.

Confidence bands for the integrated hazard of a time-homogeneous counting process

Let N be a counting process which is time-homogeneous in the sense that its conditional hazard function only depends on the covariate process Z, so N has intensity

$$\lambda_t = Y_t\, \alpha(Z_t). \tag{3.11}$$

An estimator of $\mathcal{A}(z) = \int_0^z \alpha(x)\,dx$ from i.i.d. copies (N_i, Y_i, Z_i), $i = 1, \ldots, n$ of (N, Y, Z) is given by

$$\hat{\mathcal{A}}(z) = \int_0^z \hat{A}(x)\,dx, \tag{3.12}$$

where

$$\hat{A}(x) = \int_0^1 \frac{1}{Y_r^{(n)}(s)} dN_r^{(n)}(s) \qquad \text{for } x \in I_r,$$

and $N_r^{(n)}$ is defined by (3.1) with X replaced by N. To apply our result to this special case we note that the projection $\pi: D_2 \to D[0,1]$ defined by $\pi(f)(z) = f(1,z)$ is continuous, so by the continuous mapping theorem (Billingsley, 1968, Theorem 5.1) we obtain the following consequence of Theorem 3.1. A similar result could be obtained in the case that X has continuous sample paths.

PROPOSITION 3.3. Suppose that A1-A3, B1, B2 are satisfied, $d_n^2/n \to \infty$ and $d_n = o(n^\delta)$ for some $\delta \in (1/2, 1)$. Then, for $\hat{\mathcal{A}}$ defined by (3.12),

$$\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}) \xrightarrow{\mathcal{D}} m$$

in $C[0,1]$ as $n \to \infty$, where $m = (m(z), z \in [0,1])$ is a continuous Gaussian martingale with mean zero and covariance function $\mathrm{Cov}(m(z_1), m(z_2)) = H(z_1 \wedge z_2)$, where

$$H(z) = \int_0^z \int_0^1 \frac{\alpha(x)}{f_{Z(s)Y(s)}(x,1)} ds\,dx.$$

With the help of Proposition 3.3 we now construct confidence bands for $\mathcal{A}$. Denote

$$\widehat{H}(z) = \frac{n}{d_n^2} \sum_{r=1}^{[zd_n]} \int_0^1 \frac{1}{(Y_r^{(n)}(s))^2} dN_r^{(n)}(s).$$

As a consequence of the proposition,

$$\sqrt{n}\frac{\sqrt{H(1)}}{H(\cdot) + H(1)}(\hat{\mathcal{A}}(\cdot) - \mathcal{A}(\cdot)) \xrightarrow{\mathcal{D}} W^0\left(\frac{H(\cdot)}{H(\cdot) + H(1)}\right)$$

in $C[0,1]$ as $n \to \infty$, where W^0 is a standard Brownian bridge. Now $\widehat{H}$ is a uniformly consistent estimator of H by Lemma 9. Thus we obtain the following asymptotic $100(1-\alpha)\%$ confidence band for $\mathcal{A}$:

$$\hat{\mathcal{A}}(z) \pm c_\alpha \sqrt{\frac{\widehat{H}(1)}{n}}\left(1 + \frac{\widehat{H}(z)}{\widehat{H}(1)}\right), \quad z \in [0,1],$$

where $P(\sup_{0 \le t \le 1/2} |W^0(t)| \ge c_\alpha) = \alpha$, $0 < \alpha < 1$. A table for c_α can be found in Hall and Wellner (1980).

4. Goodness-of-fit tests

4.1. Testing for independence from the covariate process

In this section we consider the problem of testing whether the covariate process Z is absent from the model, i.e. whether α is only a function of time. Let H_0 denote the null hypothesis $H_0\colon \alpha(t, z_1) = \alpha(t, z_2)$ for all $t, z_1, z_2 \in [0,1]$. Under H_0 the natural estimator of $\mathcal{A}$ is

$$\bar{\mathcal{A}}(t,z) = z\,\bar{A}(t),$$

where $\bar{A}$ is the Nelson-Aalen estimator

$$\bar{A}(t) = \int_0^t \frac{d\bar{X}^{(n)}(s)}{\bar{Y}^{(n)}(s)}$$

and

$$\bar{X}^{(n)}(t) = \sum_{i=1}^{n} \int_0^t I(Z_i(s) \in [0,1])\, Y_i(s)\, dX_i(s),$$

$$\bar{Y}^{(n)}(s) = \sum_{i=1}^{n} I(Z_i(s) \in [0,1])\, Y_i(s).$$

Define some functions g and ρ by

$$g(t,z) = \gamma(t,z,1)\, f_{Z(s)Y(s)}(z,1) \,/\, \rho^2(t),$$

$$\rho(t) = P(0 \le Z(t) \le 1,\, Y(t) = 1) = \int_0^1 f_{Z(t)Y(t)}(x,1)\, dx.$$

The following result gives the asymptotic distribution of $\hat{\mathcal{A}} - \bar{\mathcal{A}}$.

THEOREM 4.1. Under the conditions of Theorem 3.1, if H_0 holds then

$$\sqrt{n}(\hat{\mathcal{A}} - \bar{\mathcal{A}}) \xrightarrow{\mathcal{D}} m_0$$

in D_2 as $n \to \infty$, where

$$m_0(t,z) = \int_0^t \int_0^z \sqrt{h(s,x)}\, dW(s,\, x) - z \int_0^t \int_0^1 \sqrt{g(s,x)}\, dW(s,x).$$

The Kolmogorov-Smirnov type test statistic $T^{(n)} = \sqrt{n}\, \sup_{t,z} |\hat{\mathcal{A}}(t,z) - \bar{\mathcal{A}}(t,z)|$ could be used for testing H_0. Note that the continuous mapping theorem and Theorem 4.1 imply that $T^{(n)} \xrightarrow{\mathcal{D}} \sup_{t,z} |m_0(t,z)|$ as $n \to \infty$. In order to construct an asymptotic size α test of H_0, rejecting H_0 if $T^{(n)}$ is large, we first need to introduce appropriate estimators for the functions g and h under H_0. Again we shall only do this in the counting process case. Let

$$\bar{G}(t,z) = n \sum_{r=1}^{[zd_n]} \int_0^t \frac{Y_r^{(n)}(s)}{(\bar{Y}^{(n)}(s))^3}\, d\bar{X}^{(n)}(s),$$

$$\bar{H}(t,z) = \frac{n}{d_n^2} \sum_{r=1}^{[zd_n]} \int_0^t \frac{d\bar{X}^{(n)}(s)}{Y_r^{(n)}(s)\, \bar{Y}^{(n)}(s)}$$

and define

$$\bar{g}(t,z) = \frac{1}{b_n^2}\int_0^1\int_0^1 K\left(\frac{t-s}{b_n}\right)K\left(\frac{z-x}{b_n}\right)\,d\bar{G}(s,x),$$

$$\bar{h}(t,z) = \frac{1}{b_n^2}\int_0^1\int_0^1 K\left(\frac{t-s}{b_n}\right)K\left(\frac{z-x}{b_n}\right)\,d\bar{H}(s,x),$$

where K is a bounded, nonnegative kernel function with compact support, integral 1 and b_n is a bandwidth parameter, $b_n \to 0$.

The distribution of $T = \sup_{t,z}|m_0(t,z)|$ depends only on $\theta = (g,h)$ and is continuous, see Ylvisaker(1968). Let $c_\alpha(\theta)$ denote the upper α-quantile of T, so that $P_\theta\{T > c_\alpha(\theta)\} = \alpha$ for $0 < \alpha < 1$. Given the estimate $\hat{\theta}_n = (\bar{g},\bar{h})$, we may simulate the process m_0, with $\bar{g}$ and $\bar{h}$ in place of g and h respectively, to obtain an approximate critical level $c_\alpha^{(n)} = c_\alpha(\hat{\theta}_n)$. In Proposition 4.2 we show that

$$\lim_{n\to\infty} P(T^{(n)} > c_\alpha^{(n)}) = \alpha.$$

Thus, rejecting H_0 when $T^{(n)} > c_\alpha^{(n)}$ yields an asymptotic size α test for independence. In Proposition 4.3 we show that this test is consistent against all alternatives.

Proof of Theorem 4.1.
Decomposing $\bar{A}$ in a similar way to $\hat{A}$ in the proof of Theorem 3.1, we can write

$$\sqrt{n}(\hat{A} - \bar{A})(t,z) = \widehat{M}(t,z) - z\bar{M}(t) + \sqrt{n}(\hat{A}_p - \bar{A}_p)(t,z),$$

where

$$\bar{M}^{(n)}(t) = \sqrt{n}\sum_{r=1}^{d_n}\int_0^t \frac{dM_r^{(n)}(s)}{\bar{Y}^{(n)}(s)},$$

and under H_0

$$\bar{A}_p(t,z) = z\int_0^t \alpha(s)I(\bar{Y}^{(n)}(s) > 0)\,ds.$$

Putting $d_n \equiv 1$, $k \equiv 1$ in Lemma 4 of McKeague and Utikal (1987), we obtain $\sqrt{n}\|\bar{A}_p - A\|\xrightarrow{P}0$ under H_0. Also, by Lemma 1, $\sqrt{n}\|\hat{A}_p - A\|\xrightarrow{P}0$. Therefore $\sqrt{n}\|\hat{A}_p - \bar{A}_p\|\xrightarrow{P}0$ under H_0. To complete the proof it suffices to show that $\xi\xrightarrow{\mathcal{D}}m_0$, where $\xi(t,z) = \widehat{M}(t,z) - z\bar{M}(t)$. Set

$$\bar{m}(t) = \int_0^t\int_0^1 \sqrt{g(s,x)}\,dW(s,x),$$

where W is the same Brownian sheet used to define m in Theorem 3.1. Then $\bar{m}$ is a zero mean continuous Gaussian martingale with predictable variation process

$$\langle\bar{m}\rangle_t = \int_0^t\int_0^1 g(s,x)\,dx\,ds.$$

Suppose that $(\widetilde{M},\bar{M})\xrightarrow{\mathcal{D}}(m,\bar{m})$ jointly in $D_2\times D[0,1]$. Define a map $\pi_n'\colon D_2\times D[0,1]$ by $\pi_n'(f_1,f_2) = (\pi_n(f_1),f_2)$, where π_n is defined in the proof of Theorem 3.1. Then, as in that proof, $(\widehat{M},\bar{M}) = \pi_n'(\widetilde{M},\bar{M})\xrightarrow{\mathcal{D}}(m,\bar{m})$ jointly in $D_2\times D[0,1]$ and, since m and $\bar{m}$ have continuous paths, by the

continuous mapping theorem we may conclude that ξ converges weakly to the process $m(t,z) - z\bar{m}(t) = m_0(t,z)$.

It remains to show that $(\widetilde{M}, \bar{M}) \xrightarrow{\mathcal{D}} (m, \bar{m})$ jointly in $D_2 \times D[0,1]$. The process $\bar{M}$ is a martingale and $\langle \bar{M} \rangle_t \xrightarrow{P} \langle \bar{m} \rangle_t$, by Lemma 9(a). The Lindeberg condition (3.10), with $p = 1$ and $H_{1r}^{(n)}(s) = \sqrt{n}/\bar{Y}^{(n)}(s)$ can be checked as in the proof of Lemma 6. Therefore, by Rebolledo's martingale central limit theorem, $\bar{M} \xrightarrow{\mathcal{D}} \bar{m}$ in $D[0,1]$. Also, by the proof of Theorem 3.1, we have $\widetilde{M} \xrightarrow{\mathcal{D}} m$ in D_2. If we can show that the finite dimensional distributions of $(\widetilde{M}, \bar{M})$ converge to those of $(m, \bar{m})$, then $(\widetilde{M}, \bar{M}) \xrightarrow{\mathcal{D}} (m, \bar{m})$ jointly in $D_2 \times D[0,1]$ and, since m and $\bar{m}$ have continuous paths, by the continuous mapping theorem we may conclude that ξ converges weakly to the process $m(t,z) - z\bar{m}(t) = m_0(t,z)$.

To show that the finite dimensional distributions of $(\widetilde{M}, \bar{M})$ converge to those of $(m, \bar{m})$, it suffices to show that for any $0 \le z_0 < z_1 < \cdots < z_p \le 1$, $p \ge 1$,

$$((\widetilde{M}(\cdot, z_j) - \widetilde{M}(\cdot, z_{j-1}))_{j=1}^p, \bar{M}(\cdot)) \xrightarrow{\mathcal{D}} ((m(\cdot, z_j) - m(\cdot, z_{j-1}))_{j=1}^p, \bar{m}(\cdot))$$

in $D[0,1]^{p+1}$. This is done using Rebolledo's martingale central limit theorem, as in the proof of Theorem 3.1. It only remains to consider the covariation between $\widetilde{M}(\cdot, z)$ and $\bar{M}(\cdot)$. By Lemma 9(b)

$$\langle \widetilde{M}(\cdot, z), \bar{M}(\cdot) \rangle_t \xrightarrow{P} \langle m(\cdot, z), \bar{m}(\cdot) \rangle_t,$$

for each z. There are $p+1$ Lindeberg conditions to check. But these conditions have already been checked separately for the p components involving $\widetilde{M}$ and the one component involving $\bar{M}$. This completes the proof.

PROPOSITION 4.2. Suppose that X is a counting process, the assumptions of Theorem 3.1 hold, $d_n b_n^2 \to \infty$ and K is Lipschitz. Then if H_0 holds, for all $0 < \alpha < 1$

$$\lim_{n \to \infty} P(T^{(n)} > c_\alpha^{(n)}) = \alpha.$$

Proof. Let Θ denote the space of all functions of the form $\theta = (g, h)$ with g and h nonnegative bounded functions on $[0,1]^2$, and endow it with the product metric from $L^2([0,1]^2, ds\,dx) \times L^2([0,1]^2, ds\,dx)$. Let $\theta_n = (g_n, h_n)$, $n \ge 1$ be a sequence in Θ such that $\theta_n \to \theta$. Then $\sqrt{g_n} \to \sqrt{g}$ and $\sqrt{h_n} \to \sqrt{h}$ in $L^2([0,1]^2, ds\,dx)$. An argument using Doob's inequality applied twice (cf. Cairoli (1970) and Bass (1988)) shows that if $\phi \in L^2([0,1]^2, ds\,dx)$, then

$$E \sup_{t,z} \left| \int_0^t \int_0^z \phi(s,x)\, dW(s,x) \right|^2 \le 16 \int_0^1 \int_0^1 \phi^2(s,x)\, ds\, dx.$$

Applying this inequality to $\phi = \sqrt{g_n} - \sqrt{g}$ and $\phi = \sqrt{h_n} - \sqrt{h}$ gives $F_{\theta_n} \xrightarrow{\mathcal{D}} F_\theta$, where F_θ is the distribution function of T under P_θ. Let F_θ^{-1} denote the left-continuous inverse of F_θ. By Billingsley (1986, p.343) we get $c_\alpha(\theta_n) = F_{\theta_n}^{-1}(1-\alpha) \to F_\theta^{-1}(1-\alpha) = c_\alpha(\theta)$, provided F_θ^{-1} is continuous at $1-\alpha$. Now by Lemma 10(b) we have $\hat{\theta}_n \xrightarrow{P_\theta} \theta$ in the metric of Θ. Thus, using a subsequence argument, $c_\alpha(\hat{\theta}_n) \xrightarrow{P_\theta} c_\alpha(\theta)$ and

$$P_\theta(T^{(n)} > c_\alpha(\hat{\theta}_n)) \to P_\theta(T > c_\alpha(\theta)) = \alpha, \tag{4.1}$$

for all but countably many α, where we have used Slutsky's theorem and the continuity of F_θ. Since $P_\theta(T^{(n)} > c_\alpha(\hat{\theta}_n))$ is a nondecreasing function of α it follows that (4.1) holds for all $0 < \alpha < 1$, completing the proof.

PROPOSITION 4.3. Under the assumptions of Theorem 3.1, if H_0 does not hold then $T^{(n)} \xrightarrow{P} \infty$ as $n \to \infty$.

Proof. First note that if H_0 does not hold then $\|\mathcal{A} - \mathcal{A}_0\| > 0$, where

$$\mathcal{A}_0(t,z) = z \int_0^t \frac{1}{\rho(s)} \int_0^1 \alpha(s,x)\, f_{Z(s)\,Y(s)}(x,1)\, dx\, ds.$$

Also note that by Doob's inequality and Lemma 7 we have $E\|\bar{M}^{(n)}\|^2 = O(1)$, where $\bar{M}^{(n)}$ is defined in the proof of Theorem 4.1, and using similar arguments to the proof of Lemma 7

$$E\|\bar{\mathcal{A}}_p - \mathcal{A}_0\|^2 \le \sup_s E\left|\frac{\bar{\alpha}^{(n)}(s)}{\bar{Y}^{(n)}(s)} - \frac{1}{\rho(s)}\int_0^1 \alpha(s,x)\, f_{Z(s)\,Y(s)}(x,1)\, dx\right|^2 = O\Big(\frac{1}{n}\Big),$$

where

$$\bar{\mathcal{A}}_p(t,z) = z \int_0^t \frac{\bar{\alpha}^{(n)}(s)}{\bar{Y}^{(n)}(s)}\, ds \qquad \text{and} \qquad \bar{\alpha}^{(n)}(s) = \sum_{r=1}^{d_n} \alpha_r^{(n)}(s).$$

Thus, since $\sqrt{n}\|\mathcal{A} - \hat{\mathcal{A}}\| = O_P(1)$ by Theorem 3.1,

$$\sqrt{n}\|\mathcal{A} - \mathcal{A}_0\| \le \sqrt{n}\|\mathcal{A} - \hat{\mathcal{A}}\| + T^{(n)} + \|\bar{M}^{(n)}\| + \sqrt{n}\|\bar{\mathcal{A}}_p - \mathcal{A}_0\| = T^{(n)} + O_P(1).$$

This shows that $T^{(n)} \xrightarrow{P} \infty$ if H_0 does not hold.

REMARK. The above test for independence can be modified to provide a goodness-of-fit test for Aalen's multiplicative intensity model. Now H_0 is the null hypothesis H_0: there exists a function $\alpha_0\colon [0,1] \to \mathbf{R}$ such that $\alpha(t,z) = \alpha_0(t)\, z$ for all t, $z \in [0,1]$. Under this H_0, the natural estimator of $\mathcal{A}$ is $\bar{\mathcal{A}}(t,z) = \frac{1}{2} z^2 \bar{A}(t)$, where $\bar{A}$ is the Nelson-Aalen estimator as before, except that

$$\bar{Y}^{(n)}(s) = \sum_{i=1}^n I(Z_i(s) \in [0,1])\, Y_i(s)\, Z_i(s).$$

The only changes to Theorem 4.1 are that

$$\rho(t) = \int_0^1 x\, f_{Z(t)\,Y(t)}(x,1)\, dx$$

and

$$m_0(t,z) = \int_0^t \int_0^z \sqrt{h(s,x)}\, dW(s,x) - \frac{1}{2} z^2 \int_0^t \int_0^1 \sqrt{g(s,x)}\, dW(s,x).$$

4.2. *Testing for time-homogeneity*

We want to derive a test for the hypothesis $H_0\colon \alpha(t_1, z) = \alpha(t_2, z)$ for all t_1, t_2, $z \in [0,1]$, i.e. α is only a function of the covariate. One possible application of such a test would be in testing whether a pure jump process on a finite state space is a Markov renewal process, see McKeague and Utikal (1987, Section 1). The natural estimator for $\mathcal{A}$ under H_0 is $\mathcal{A}^*(t,z) = t\hat{\mathcal{A}}(1,z)$. In order to test H_0 we could use the test statistic $S^{(n)} = \sqrt{n}\sup_{t,z}|\hat{\mathcal{A}}(t,z) - \mathcal{A}^*(t,z)|$. As in Section 4.1, once we know the asymptotic distribution of $\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}^*)$ we can derive an asymptotic size α test for H_0 based on $S^{(n)}$. This test can be shown to be consistent using a proof similar to that of Proposition 4.3. The asymptotic distribution of $\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}^*)$ is given by the following theorem.

THEOREM 4.4. Under the conditions of Theorem 3.1, if H_0 holds then

$$\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}^*) \xrightarrow{\mathcal{D}} m_1$$

in D_2 as $n \to \infty$, where

$$m_1(t,z) = \int_0^t \int_0^z \sqrt{h(s,x)}\, dW(s,x) - t\int_0^1 \int_0^z \sqrt{h(s,x)}\, dW(s,x).$$

Proof. Note that $\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}^*) = \pi(\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}))$, where $\pi\colon D_2 \to D_2$ defined by $\pi(f)(t,z) = f(t,z) - t\,f(1,z)$ is continuous. The result follows immediately, using Theorem 3.1 and the continuous mapping theorem.

4.3. *Testing for proportionality*

Thomas (1983) introduced the model $\alpha(t,z) = \alpha_1(t)\,\alpha_2(z)$ for the conditional hazard function in the survival analysis context, where $\alpha_j\colon [0,1] \to \mathbf{R}$, $j = 1,2$ are unknown functions. This model is a generalization of Cox's proportional hazards model to allow for arbitrary covariate dependence while keeping the proportional hazards form. In this section we introduce a goodness-of-fit test for Thomas' model. Note that this is not the same as a goodness-of-fit test for Cox's proportional hazards model. However, Cox's model can be treated in a similar fashion, see McKeague and Utikal (1988).

Let H_0 denote the null hypothesis H_0: there exist functions $\alpha_j\colon [0,1] \to \mathbf{R}$, $j = 1,2$ such that $\alpha(t,z) = \alpha_1(t)\,\alpha_2(z)$ for all $t, z \in [0,1]$. In order that α_1 and α_2 are identifiable we impose the condition $A_1(1) = 1$ under H_0, where $A_1(t) = \int_0^t \alpha_1(s)\,ds$. Equivalently, we could impose the condition $A_2(1) = 1$, where $A_2(z) = \int_0^z \alpha_2(x)\,dx$. A reasonable estimator for $\mathcal{A}$ under H_0 is

$$\mathcal{A}^\dagger(t,z) = \hat{A}_1(t)\,\hat{A}_2(z),$$

where

$$\hat{A}_1(t) = \frac{\hat{\mathcal{A}}(t,1)}{\hat{\mathcal{A}}(1,1)} \quad \text{(with } 1/0 \equiv 0) \qquad \text{and} \qquad \hat{A}_2(z) = \hat{\mathcal{A}}(1,z).$$

In order to test H_0 we could use the test statistic $U^{(n)} = \sqrt{n}\sup_{t,z}|\hat{\mathcal{A}}(t,z) - \mathcal{A}^\dagger(t,z)|$. As before, once we know the asymptotic distribution of $\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}^\dagger)$ we can derive an asymptotic size α test for H_0 based on $U^{(n)}$. This test is an omnibus goodness-of-fit test for proportionality in that it is consistent against any alternative.

THEOREM 4.5. Under the conditions of Theorem 3.1, if H_0 holds and $A_2(1) \neq 0$, then

$$\sqrt{n}(\hat{\mathcal{A}} - \mathcal{A}^{\dagger}) \xrightarrow{\mathcal{D}} m_2$$

in D_2 as $n \to \infty$, where

$$m_2(t,z) = \int_0^t \int_0^z \sqrt{h(s,x)}\, dW(s,x) - \beta\, A_2(z) \int_0^t \int_0^1 \sqrt{h(s,x)}\, dW(s,x)$$

$$-\beta\, A_1(t) \int_0^1 \int_0^z \sqrt{h(s,x)}\, dW(s,x) + \beta\, A_1(t)\, A_2(z) \int_0^1 \int_0^1 \sqrt{h(s,x)}\, dW(s,x)$$

and $\beta = 1/A_2(1)$.

Proof. The result follows readily from Theorem 3.1, using the continuous mapping theorem (cf. the proof of Theorem 4.4) and the identities

$$(\hat{\mathcal{A}} - \mathcal{A}^{\dagger})(t,z) = (\hat{\mathcal{A}} - \mathcal{A})(t,z) - \frac{1}{\mathcal{A}(1,1)}[\hat{\mathcal{A}}(t,1)\,\hat{\mathcal{A}}(1,z) - \mathcal{A}(t,1)\,\mathcal{A}(1,z)]$$
$$+ \hat{\mathcal{A}}(t,1)\,\hat{\mathcal{A}}(1,z)\Big[\frac{1}{\mathcal{A}(1,1)} - \frac{1}{\hat{\mathcal{A}}(1,1)}\Big]$$

$$= (\hat{\mathcal{A}} - \mathcal{A})(t,z) - \frac{1}{\mathcal{A}(1,1)}[\,(\hat{\mathcal{A}}(t,1) - \mathcal{A}(t,1))\,(\hat{\mathcal{A}}(1,z) - \mathcal{A}(1,z)\,) + \mathcal{A}(1,z)(\hat{\mathcal{A}}(t,1) - \mathcal{A}(t,1))$$
$$+\mathcal{A}(t,1)\,(\hat{\mathcal{A}}(1,z) - \mathcal{A}(1,z)] + \frac{\mathcal{A}(t,1)\,\mathcal{A}(1,z)}{(\mathcal{A}(1,1)\,)^2}(\hat{\mathcal{A}}(1,1) - \mathcal{A}(1,1)\,)$$
$$+ \frac{\hat{\mathcal{A}}(1,1) - \mathcal{A}(1,1)}{\mathcal{A}(1,1)}\Big[\frac{\hat{\mathcal{A}}(t,1)\hat{\mathcal{A}}(1,z)}{\hat{\mathcal{A}}(1,1)} - \frac{\mathcal{A}(t,1)\mathcal{A}(1,z)}{\mathcal{A}(1,1)}\Big].$$

REMARK. Under H_0, we have from Theorem 3.1 that $\hat{A}_1$ and $\hat{A}_2$ are uniformly consistent estimators of A_1 and A_2, respectively, and

$$\sqrt{n}(\hat{A}_2 - A_2) \xrightarrow{\mathcal{D}} m_3 \tag{4.9}$$

in $D[0,1]$ as $n \to \infty$, where m_3 is a continuous Gaussian martingale with covariance function

$$\mathrm{Cov}(m_3(z_1), m_3(z_2)) = \int_0^{z_1 \wedge z_2} \int_0^1 h(s,x)\, ds\, dx.$$

This could be used to obtain confidence bands for A_2 under Thomas' model, by transforming m_3 to Brownian bridge (cf. the discussion following Proposition 3.2). An analogous result can be obtained for $\hat{A}_1$.

5. Technical Lemmas

In this section we make frequent use of lemmas from McKeague and Utikal (1987). We apply those lemmas by changing w_n to d_n^{-1}, by taking $I_z = I_r$ when $z \in I_r$, by changing $Y^{(n)}(s,z)$ to $Y_r^{(n)}(s)$, $\gamma^{(n)}(s)$ to $\gamma_r^{(n)}(s)$ and $J^{(n)}(s,z)$ to $J_r^{(n)}(s)$, where $J_r^{(n)}(s) = I\{Y_r^{(n)}(s) \neq 0\}$. Since $\cup_{1 \le r \le d_n} I_r = [0,1]$ for all n, the set C in A1-A3, B1, B2 of McKeague and Utikal (1987) can be chosen as $[0,1]$. Results quoted from McKeague and Utikal (1987) will be referred to as Lemma A.1, etc..

LEMMA 1. Suppose that A1, A2 and B2 hold and

$$d_n = o(n^{\delta}) \text{ for some } \delta \in (1/2, 1). \tag{5.1}$$

Then

$$E\,\|\hat{\mathcal{A}}_p - \mathcal{A}\,\| = O(d_n^{-1}).$$

Proof. From (3.8) we have

$$\begin{aligned} E\|\hat{\mathcal{A}}_p - \mathcal{A}\,\| &= E \sup_{t,z} \Big| \sum_{r=1}^{d_n} \int_0^z \int_0^t \frac{\alpha_r^{(n)}(s)}{Y_r^{(n)}(s)}\, ds\, I(x \in I_r)\, dx - \int_0^z \int_0^t \alpha(s,x)\, ds\, dx \Big| \\ &\le \sup_{s,r,x \in I_r} E \Big| \frac{\alpha_r^{(n)}(s)}{Y_r^{(n)}(s)} - \alpha(s,x) \Big| \le I_1 + I_2, \end{aligned}$$

where

$$\begin{aligned} I_1 &= \sup_{s,r,x \in I_r} E \Big| \frac{\alpha_r^{(n)}(s)}{Y_r^{(n)}(s)} - J_r^{(n)}(s)\, \alpha(s,x) \Big|, \\ I_2 &= \sup_{r,s,x} |\alpha(s,x)|\, E\,(1 - J_r^{(n)}(s)\,). \end{aligned}$$

By Lemma A.5 we have $I_1 = O(d_n^{-1})$. Next, by Lemma A.4 we have $I_2 = O(\exp\{-n\,K\,d_n^{-1}\})$, where $K = \inf_{t,z} f_{Z(t)\,Y(t)}(z,1)$. Thus, since $\exp\{-n\,K\,d_n^{-1}\} = O(d_n/n)^k$ for all nonnegative integers k and $(d_n/n)^k = O(1/d_n)$ for k sufficiently large by (5.1), we have $I_2 = O(d_n^{-1})$. This completes the proof.

Proof of tightness

First note that

$$d\langle M_r^{(n)} \rangle_s = \sum_{i=1}^{n} I\{Z_i(s) \in I_r\} Y_i(s) d\langle M_i \rangle_s \tag{5.2}$$

and, since $M_r^{(n)}$, $r = 1, \ldots, d_n$ are orthogonal martingales,

$$d\langle \widetilde{M}^{(n)}(\cdot, z) \rangle_s = \frac{n}{d_n^2} \sum_{r=1}^{[zd_n]} \frac{1}{(Y_r^{(n)}(s)\,)^2} d\langle M_r^{(n)} \rangle_s. \tag{5.3}$$

Tightness of $\{\widetilde{M}^{(n)}, n \geq 1\}$ in D_2 will be shown by establishing a product moment condition on the increments of $\widetilde{M}^{(n)}$ over the grid $T^{(n)} = [0,1] \times \{0, 1/d_n, 2/d_n, \ldots, 1\}$.

For fixed $0 \leq s \leq 1$, $0 \leq x < y \leq 1$, define the martingale

$$M_1(t) = \widetilde{M}^{(n)}((s,t] \times (x,y]), \quad t \geq s,$$

and denote $m_1 = M_1^2 - \langle M_1 \rangle$.

LEMMA 2. Suppose that A1, A2 hold, $d_n = o(n)$ and X is either a continuous process or a counting process. Then there exists a positive constant K such that for all $n \geq 1$, (s,x), $(t,y) \in T^{(n)}$

$$E\langle M_1 \rangle_t^2 \leq K\,(t-s)^2(y-x)^2, \tag{5.4}$$

$$E\, m_1^2(t) \leq K\,(t-s)(y-x)^2. \tag{5.5}$$

Proof. By (5.3)

$$E\langle M_1 \rangle_t^2 = \frac{n^2}{d_n^4} \sum_{r_1,\, r_2 = [xd_n]+1}^{[yd_n]} E \int_s^t \frac{1}{(Y_{r_1}^{(n)}(v))^2} d\langle M_{r_1}^{(n)} \rangle_v \int_s^t \frac{1}{(Y_{r_2}^{(n)}(v))^2} d\langle M_{r_2}^{(n)} \rangle_v.$$

By (2.2) there exists a positive constant K_1 such that $d\langle M \rangle_s / ds \leq K_1$ for all $0 \leq s \leq 1$. Therefore, by (5.2), $d\langle M_r^{(n)} \rangle_s / ds \leq K_1 Y_r^{(n)}(s)$. Applying Fubini's theorem we obtain

$$\begin{aligned} &E \int_s^t \frac{1}{(Y_{r_1}^{(n)}(v))^2} d\langle M_{r_1}^{(n)} \rangle_v \int_s^t \frac{1}{(Y_{r_2}^{(n)}(v))^2} d\langle M_{r_2}^{(n)} \rangle_v \\ &\leq K_1^2 \int_s^t \int_s^t E \frac{1}{Y_{r_1}^{(n)}(v_1)\, Y_{r_2}^{(n)}(v_2)} dv_1 dv_2. \end{aligned}$$

But by Lemma A.3

$$\sup_{r,s} E\Big[\frac{1}{Y_r^{(n)}(s)}\Big]^2 = O\Big(\frac{d_n}{n}\Big)^2.$$

This proves (5.4). Now we turn to the proof of (5.5). First we need to obtain an explicit expression for m_1. Integration by parts gives

$$M_1^2(t) = 2\int_s^t M_1(v-)\, dM_1(v) + [M_1]_t.$$

In the case that X has continuous sample paths $[M_1] = \langle M_1 \rangle$. In the counting process case

$$[M_1]_t = \sum_{s<v\leq t} (\Delta M_1(v))^2,$$

where $\Delta M_1(v) = M_1(v) - M_1(v-)$ is the jump in M_1 at time v, so from (3.6), (3.9)

$$\begin{aligned} [M_1]_t &= \frac{n}{d_n^2} \sum_{r=[xd_n]+1}^{[yd_n]} \int_s^t \frac{1}{(Y_r^{(n)}(v))^2} dX_r^{(n)}(v) \\ &= \langle M_1 \rangle_t + \eta_t, \end{aligned}$$

where

$$\eta_t = \frac{n}{d_n^2} \sum_{r=[xd_n]+1}^{[yd_n]} \int_s^t \frac{1}{(Y_r^{(n)}(v))^2} dM_r^{(n)}(v).$$

Thus

$$m_1(t) = 2\int_s^t M_1(v-)\, dM_1(v) + \eta_t,$$

where in the continuous sample path case η_t is zero. From this expression we get

$$Em_1^2(t) \leq 8\, E\int_s^t M_1^2(v-)\, d\langle M_1\rangle_v + 2\, E\langle\eta\rangle_t. \tag{5.6}$$

In order to obtain an upper bound on the first term on the r.h.s. of (5.6) we shall use the Burkholder-Davis-Gundy inequality (see Dellacherie and Meyer, 1982, p.287)

$$E \sup_{v\in[s,t]} M_1^4(v) \leq K\, E[M_1]_t^2, \tag{5.7}$$

where here, and in what follows, K is a generic positive constant which is independent of n. Then, by orthogonality of the martingales $M_r^{(n)}$, $r = 1, \ldots, d_n$,

$$\begin{aligned}
E\int_s^t M_1^2(v-)\, d\langle M_1\rangle_v &= \frac{n}{d_n^2} \sum_{r=[xd_n]+1}^{[yd_n]} E\int_s^t \frac{M_1^2(v-)}{(Y_r^{(n)}(v))^2} d\langle M_r^{(n)}\rangle_v \\
&\leq \frac{n}{d_n^2} d_n\, (y-x)(t-s)\, K \sup_{r,v} E\left(\frac{M_1^2(v-)}{Y_r^{(n)}(v)}\right) \\
&\leq \frac{n}{d_n} K\, (y-x)(t-s)(E[M_1]_t^2)^{\frac{1}{2}} (\sup_{r,v} E\, (1/Y_r^{(n)}(v)\,)^2)^{\frac{1}{2}} \\
&\quad \text{(by (5.7) and the Cauchy – Schwarz inequality)} \\
&\leq K\, (y-x)(t-s)(E\, [M_1]_t^2)^{\frac{1}{2}},
\end{aligned} \tag{5.8}$$

by Lemma A.3. Now $[M_1] = \langle M_1\rangle + \eta$, so

$$E\, [M_1]_t^2 \leq 2\, E\langle M_1\rangle_t^2 + 2\, E\langle\eta\rangle_t \leq K\, (t-s)^2(x-y)^2 + 2\, E\langle\eta\rangle_t \tag{5.9}$$

by (5.4). Also

$$\begin{aligned}
E\langle\eta\rangle_t &= \frac{n}{d_n^2} \sum_{r=[xd_n]+1}^{[yd_n]} E\int_s^t \frac{1}{(Y_r^{(n)}(v))^4} d\langle M_r^{(n)}\rangle_v \\
&\leq \frac{n}{d_n^2} d_n\, (y-x)(t-s)\, K \sup_{r,v} E\left(\frac{1}{Y_r^{(n)}(v)}\right)^3 \\
&\leq K \frac{d_n}{n} \frac{(y-x)}{d_n}(t-s) \quad \text{(by Lemma A.3)} \\
&\leq K\, (y-x)^2(t-s)
\end{aligned} \tag{5.10}$$

since $d_n = o(n)$ and $y - x \geq 1/d_n$ if $x \neq y$. The desired inequality is now obtained directly from (5.6), (5.8)-(5.10).

LEMMA 3 (Tightness). Suppose that A1, A2 hold, $d_n = o(n)$ and X is either a continuous process or a counting process. Then $\{\widetilde{M}^{(n)},\ n \geq 1\}$ is tight in D_2.

Proof. Consider the following increments of $\widetilde{M}^{(n)}$ over neighbouring rectangles in $[0,1]^2$. Define M_1 as before and

$$M_2(t) = \widetilde{M}^{(n)}(\,(s,t]\,) \times (y,z]\,),$$
$$M_3(y) = \widetilde{M}^{(n)}(\,(t,u]\,) \times (x,y]\,),$$

where $0 \leq s < t < u \leq 1$, $0 \leq x < y < z \leq 1$. Suppose that the corner points of the rectangles belong to $T^{(n)}$. Also, denote $m_i = M_i^2 - \langle M_i \rangle$, $i = 1,\ 2,\ 3$. From the representation of m_1 in the proof of Lemma 2 it can be seen that m_1 and m_2 are orthogonal martingales. Thus, using the Cauchy-Schwarz inequality and Lemma 2, we get

$$\begin{aligned} E\, M_1^2(t)\, M_2^2(t) &= E\langle M_1\rangle_t \langle M_2\rangle_t + E\, m_1(t)\langle M_2\rangle_t + E\, m_2(t)\langle M_1\rangle_t + E\, m_1(t)\, m_2(t) \\ &\leq (t-s)^{\frac{3}{2}}(y-x)(z-y). \end{aligned} \tag{5.11}$$

Next, by the martingale property of m_3, we have

$$\begin{aligned} E\, M_1^2(t)\, M_3^2(u) &= E(M_1^2(t)\, E(M_3^2(u)|\mathcal{F}_t)\,) = E\,(M_1^2(t)\langle M_3\rangle_u) \\ &= E\, m_1(t)\langle M_3\rangle_u + E\langle M_1\rangle_t \langle M_3\rangle_u, \end{aligned}$$

so that, again using the Cauchy-Schwarz inequality and Lemma 2, we obtain

$$E\, M_1^2(t)\, M_3^2(u) \leq K\,(y-x)^2(t-s)^{\frac{1}{2}}(u-t). \tag{5.12}$$

The inequalities (5.11) and (5.12) imply that "condition $(\beta,\ \gamma)$" of Bickel and Wichura (1971, p.1658) is satisfied with $\beta = 3/2$, $\gamma = 4$, for rectangles whose corner points lie in $T^{(n)}$. Clearly $T^{(n)}$ becomes dense in $[0,1]^2$ as n grows large. Moreover, $\widetilde{M}^{(n)}(t,z)$ is constant as a function of z over each interval $I_r^{(n)} = [(r-1)/d_n,\ r/d_n)$, $r = 1,\ldots,d_n$, so the modulus of continuity $\omega_\delta''(\widetilde{M}^{(n)})$ defined in Bickel and Wichura can be computed using $T^{(n)}$ instead of $[0,1]^2$. Tightness of $\{\widetilde{M}^{(n)},\ n \geq 1\}$ now follows from the remarks following Theorem 3 of Bickel and Wichura (1971, p.1665).

Convergence of finite dimensional distributions

Recall the notation $H(t,z) = \int_0^t \int_0^z h(s,x)\, dx\, ds$.

LEMMA 4. Suppose that A1-A3, B1 hold and $d_n = o(n)$. Then

$$\sup_{t,z} |\langle \widetilde{M}^{(n)}(\cdot,z)\rangle_t - H(t,z)\,| \xrightarrow{L^1} 0.$$

Proof. From (2.2) and (5.2) we have $d\langle M_r^{(n)}\rangle_s = \sum_{i=1}^n I\{Z_i(s) \in I_r\}\, Y_i(s)\, \gamma(s, Z_i(s), 1\,)\, ds$. By continuity of $\gamma(\cdot,\cdot,1)$

$$I\{Z_i(s) \in I_r\}\, \gamma(s, Z_i(s), 1) = I\{Z_i(s) \in I_r\}\, (\gamma(s, x_r^{(n)}, 1) + o(1)\,) \tag{5.13}$$

and $d\langle M_r^{(n)}\rangle_s = Y_r^{(n)}(s)\, (\gamma(s, x_r^{(n)}, 1) + o(1)\,)\, ds$ for arbitrary $x_r^{(n)} \in I_r$ uniformly in $r = 1,\ldots,d_n$ and $s \in [0,1]$. Therefore by (5.3)

$$\langle \widetilde{M}^{(n)}(\cdot, z)\rangle_t = \frac{1}{d_n}\sum_{r=1}^{[zd_n]}\int_0^t \frac{n}{d_n}\frac{1}{Y_r^{(n)}(s)}(\gamma(s, x_r^{(n)}, 1) + o(1))\, ds.$$

Thus, since h is continuous,

$$\begin{aligned}
&E \sup_{t,z} |\langle \widetilde{M}^{(n)}(\cdot, z)\rangle_t - H(t, z)\,| \\
&\leq \frac{1}{d_n}\sum_{r=1}^{d_n} E\int_0^1 \Big|\frac{n}{d_n}\frac{1}{Y_r^{(n)}(s)}(\gamma(s, x_r^{(n)}, 1) + o(1)) - h(s, x_r^{(n)})\Big|\, ds + o(1) \\
&\leq I_1 + I_2 + o(1),
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \sup_{s,r} E\Big[\frac{n}{d_n}\frac{1}{Y_r^{(n)}(s)}\Big]\, o(1), \\
I_2 &= \sup_{s,r} E\Big|\frac{n}{d_n}\frac{\gamma(s, x_r^{(n)}, 1)}{Y_r^{(n)}(s)} - h(s, x_r^{(n)})\Big|.
\end{aligned}$$

It follows from Lemma A.3 that $I_1 \to 0$. From (5.13) we have that $\gamma_r^{(n)}(s) = Y_r^{(n)}(s)\,(\gamma(s, x_r^{(n)}, 1) + o(1))$, where $\gamma_r^{(n)}(s) = \sum_{i=1}^n I\{Z_i(s) \in I_r\}\, Y_i(s)\,\gamma(s, Z_i(s), 1)$. Therefore

$$\frac{\gamma(s, x_r^{(n)}, 1)}{Y_r^{(n)}(s)} = \frac{\gamma_r^{(n)}(s)}{(Y_r^{(n)}(s))^2} + \frac{1}{Y_r^{(n)}(s)} o(1).$$

Application of Lemma A.6 yields $I_2 \to 0$. This proves the lemma.

LEMMA 5. Suppose A1, A2 hold and $d_n = o(n)$. Then, if X is a counting process, the Lindeberg condition (3.10) is satisfied.

Proof. By (2.2) and (5.2) it suffices to show that

$$I_1 = \frac{n}{d_n^2}\sum_{r=1}^{d_n} E\int_0^1 \frac{1}{Y_r^{(n)}(s)} I\Big\{\frac{\sqrt{n}}{d_n}\frac{1}{Y_r^{(n)}(s)} > \epsilon\Big\}\, ds \to 0.$$

But, by the Cauchy-Schwarz inequality, Lemma A.3 and Chebychev's inequality

$$\begin{aligned}
I_1 &\leq \frac{n}{d_n}\sup_{r,s}\Big\{E\Big[\frac{1}{Y_r^{(n)}(s)}\Big]^2 P\Big[\frac{\sqrt{n}}{d_n}\frac{1}{Y_r^{(n)}(s)} > \epsilon\Big]\Big\}^{1/2} \\
&\leq \frac{n}{d_n}\Big\{O\Big(\frac{d_n}{n}\Big)^2 \frac{n}{d_n^2} O\Big(\frac{d_n}{n}\Big)^2\Big\}^{\frac{1}{2}} = O\Big(\frac{1}{\sqrt{n}}\Big),
\end{aligned}$$

which proves the lemma.

We shall make use of the following notation:

$$\begin{aligned}
\gamma_r^{(n)}(s) &= \frac{d\langle M_r^{(n)}\rangle_s}{ds} = \sum_{i=1}^n I\{Z_i(s) \in I_r\} Y_i(s)\,\gamma(s, Z_i(s), 1), \\
\bar{\gamma}^{(n)}(s) &= \sum_{r=1}^{d_n} \gamma_r^{(n)}(s).
\end{aligned}$$

LEMMA 6. Suppose that A1-A3, B1 hold and $d_n = o(n)$. Then

$$\sup_s E\Big[n\frac{\bar{\gamma}^{(n)}(s)}{(\bar{Y}^{(n)}(s))^2} - \int_0^1 g(s,x)\,dx\Big]^2 = O\Big(\frac{1}{n}\Big).$$

Proof. By the Cauchy-Schwarz inequality and Lemma A.3 (with $w_n = 1$)

$$\begin{aligned} &E\Big[n\frac{\bar{\gamma}^{(n)}(s)}{(\bar{Y}^{(n)}(s))^2} - \int_0^1 g(s,x)\,dx\Big]^2 \\ &\le \Big\{E\Big[\frac{n}{\bar{Y}^{(n)}(s)}\Big]^8\Big\}^{1/2}\Big\{E\Big[\frac{1}{n}\bar{\gamma}^{(n)}(s) - \Big(\frac{\bar{Y}^{(n)}(s)}{n\,\rho(s)}\Big)^2\int_0^1 \rho^2(s)g(s,x)\,dx\Big]^4\Big\}^{1/2} \\ &\le K\{I_1(s) + I_2(s)\}^{1/2}, \end{aligned}$$

where

$$I_1(s) = E\Big[\frac{1}{n}\bar{\gamma}^{(n)}(s) - \int_0^1 \rho^2(s)g(s,x)\,dx\Big]^4,$$

$$I_2(s) = E\Big[\rho(s) - \frac{\bar{Y}^{(n)}(s)}{n}\Big]^4.$$

Now $\bar{\gamma}^{(n)}(s)$ is a sum of i.i.d. r.v.'s, each of which is uniformly bounded in s and has expectation $\int_0^1 \rho^2(s)g(s,x)\,dx$. Thus $\sup_s I_1(s) = O(1/n^2)$. Similarly $\sup_s I_2(s) = O(1/n^2)$. This completes the proof.

LEMMA 7. Suppose that A1-A3 hold and $d_n = o(n)$. Then

$$\sup_{s,z} E\Big|\frac{n}{d_n\bar{Y}^{(n)}(s)}\sum_{r=1}^{[zd_n]}\frac{\gamma_r^{(n)}(s)}{Y_r^{(n)}(s)} - \frac{1}{\rho(s)}\int_0^z \gamma(s,x,1)\,dx\Big| \to 0.$$

Proof. Using the Cauchy-Schwarz inequality and Lemma A.3, as in the proof of Lemma 6, we see that it suffices to show that $I_1 \to 0$ and $I_2 \to 0$, where

$$I_1 = \sup_{s,z} E\Big[\frac{1}{d_n}\sum_{r=1}^{[zd_n]}\frac{\gamma_r^{(n)}(s)}{Y_r^{(n)}(s)} - \int_0^z \gamma(s,x,1)\,dx\Big]^2,$$

$$I_2 = \sup_s E\,[\frac{1}{n}\bar{Y}^{(n)}(s) - \rho(s)]^2.$$

Since $\gamma(\cdot,\cdot,1)$ is continuous, $\int_0^z \gamma(s,x,1)\,dx = d_n^{-1}\sum_{r=1}^{[zd_n]}\gamma(s,x_r,1) + o(1)$ uniformly in s and z, where $x_r \in I_r$ is arbitrary. Thus

$$I_1 \le \sup_{r,s} E\Big|\frac{\gamma_r^{(n)}(s)}{Y_r^{(n)}(s)} - \gamma(s,x_r,1)\Big| + o(1)$$

which tends to zero by Lemmas A.4, A.5 (with α replaced by γ). Finally, $I_2 \to 0$ by the proof of Lemma 6.

LEMMA 8. Suppose that A1, A2 hold and $d_n = o(n)$. Then

(a) $\langle \bar{M}\rangle_t \xrightarrow{P} \int_0^t \int_0^1 g(s,x)\,dx\,ds$;

(b) $\langle \widetilde{M}(\cdot,z), \bar{M}(\cdot)\rangle_t \xrightarrow{P} \int_0^t \frac{1}{\rho(s)} \int_0^z \gamma(s,x,1)\,dx\,ds$.

Proof. Using the orthogonality of the martingales $M_r^{(n)}$, $r = 1,\ldots,d_n$, we have

$$\langle \bar{M}\rangle_t = n \int_0^t \frac{\bar{\gamma}^{(n)}(s)}{(\bar{Y}^{(n)}(s))^2}\,ds,$$

$$\langle \widetilde{M}(\cdot,z), \bar{M}(\cdot)\rangle_t = \frac{n}{d_n} \sum_{r=1}^{[zd_n]} \int_0^t \frac{\gamma_r^{(n)}(s)}{Y_r^{(n)}(s)\,\bar{Y}^{(n)}(s)}\,ds.$$

Parts (a) and (b) the follow immediately from Lemmas 6 and 7, respectively.

Estimation of H, h and g

LEMMA 9. Suppose that A1-A3, B1 hold, X is a counting process and $d_n = o(n)$. Then

$$\sup_{t,z} |\,\hat{H}(t,z) - H(t,z)\,| \xrightarrow{L^1} 0.$$

Proof. From (3.4)-(3.6) and (5.3) we obtain $\hat{H} - H = I_1 + I_2$, where

$$I_1(t,z) = \frac{n}{d_n^2} \sum_{r=1}^{[zd_n]} \int_0^t \frac{1}{(Y_r^{(n)}(s))^2}\,dM_r^{(n)}(s),$$

$$I_2(t,z) = \langle \widetilde{M}^{(n)}(\cdot,z)\rangle_t - H(t,z).$$

From Lemma 4 we have that $\sup_{t,z} |I_2(t,z)| \xrightarrow{L^1} 0$. By Doob's inequality, (5.2), (2.2), (2.3) and Lemma A.3 we get

$$\begin{aligned} E \sup_{t,z} I_1^2(t,z) &\le 4\,\frac{n^2}{d_n^4} \sum_{r=1}^{d_n} E \int_0^1 \frac{1}{(Y_r^{(n)}(s))^4}\,d\langle M_r^{(n)}\rangle_s \\ &\le 4\,\frac{n^2}{d_n^3} \sup_{s,z} \alpha(s,z) \sup_{r,s} E \frac{1}{(Y_r^{(n)}(s))^3} = O\Big(\frac{n^2}{d_n^3}\Big)\,O\Big(\frac{d_n}{n}\Big)^3 = O\Big(\frac{1}{n}\Big). \end{aligned}$$

LEMMA 10. Suppose that X is a counting process, the assumptions of Theorem 3.1 hold, $d_n b_n^2 \to \infty$ and K is Lipschitz. Then

(a) $E \int_0^1 \int_0^1 |\hat{h}(t,z) - h(t,z)|^2\,dt\,dz \to 0$;

(b) under the null hypothesis H_0 of Section 4.1, $E \int_0^1 \int_0^1 |\bar{h}(t,z) - h(t,z)|^2\,dt\,dz \to 0$ and $E \int_0^1 \int_0^1 |\bar{g}(t,z) - g(t,z)|^2\,dt\,dz \to 0$.

Proof. We shall prove (a); the proof of (b) is similar. From (3.6) and the definition of $\hat{h}$, we can write

$$h - \hat{h} = (h - \tilde{h}) + (\tilde{h} - h^\dagger) + (h^\dagger - h^*) - R, \tag{5.14}$$

where

$$\tilde{h}(t,z) = \frac{1}{b_n^2}\int_0^1\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,K\Big(\frac{z-x}{b_n}\Big)\,h(s,x)\,ds\,dx,$$

$$h^{\dagger}(t,z) = \frac{1}{b_n^2 d_n}\sum_{r=1}^{d_n} K\Big(\frac{z-x_r}{b_n}\Big)\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,h(s,x)\,ds,$$

$$h^{*}(t,z) = \frac{1}{b_n^2 d_n}\sum_{r=1}^{d_n} K\Big(\frac{z-x_r}{b_n}\Big)\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\frac{n}{d_n}\frac{\alpha_r^{(n)}(s)}{(Y_r^{(n)}(s))^2}\,ds,$$

$$R(t,z) = \frac{n}{b_n^2 d_n^2}\sum_{r=1}^{d_n} K\Big(\frac{z-x_r}{b_n}\Big)\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\frac{1}{(Y_r^{(n)}(s))^2}\,dM_r^{(n)}(s),$$

and $x_r = r/d_n$. Now let us treat each term in (5.14) separately. First, since h is continuous,

$$\int_0^1\int_0^1 (h(t,z) - \tilde{h}(t,z))^2\,dt\,dz \to 0.$$

Secondly, since h is continuous and K is Lipschitz,

$$\begin{aligned}
\sup_{t,z}|\tilde{h}(t,z) - h^{\dagger}(t,z)| &\le \frac{1}{b_n}\sup_{s,z}\Big|\int_0^1 K\Big(\frac{z-x}{b_n}\Big)\,h(s,x)\,dx - \frac{1}{d_n}\sum_{r=1}^{d_n} K\Big(\frac{z-x_r}{b_n}\Big)\,h(s,x_r)\Big| \\
&\le \frac{1}{b_n}\sup_{s,z}\Big\{\sum_{r=1}^{d_n}\int_{I_r}\Big|K\Big(\frac{z-x}{b_n}\Big) - K\Big(\frac{z-x_r}{b_n}\Big)\Big|\,dx\,h(s,x_r)\Big\} + o(1) \\
&\le \frac{1}{b_n}O\Big(\frac{1}{b_n d_n}\Big) + o(1) \to 0.
\end{aligned}$$

Thirdly, using Lemma A.6 and the assumption that K has compact support,

$$\sup_{t,z} E|h^{\dagger}(t,z) - h^{*}(t,z)|^2 \le$$

$$\frac{1}{b_n^4 d_n^2}\sup_t\Big[\int_0^1 K\Big(\frac{t-s}{b_n}\Big)\,ds\Big]^2 \sup_z\Big[\sum_{r=1}^{d_n} K\Big(\frac{z-x_r}{b_n}\Big)\Big]^2 \sup_{s,r} E\Big|h(s,x_r) - \frac{n}{d_n}\frac{\alpha_r^{(n)}(s)}{(Y_r^{(n)}(s))^2}\Big|^2$$

$$= \frac{1}{b_n^4 d_n^2}O(b_n)^2 O(b_n d_n)^2 o(1) \to 0.$$

Finally, using Lemma A.3,

$$\begin{aligned}
\sup_{z,t} E\,|R(t,z)|^2 &= \frac{n^2}{b_n^4 d_n^2}\sup_{t,z}\Big\{\sum_{r=1}^{d_n} K^2\Big(\frac{z-x_r}{b_n}\Big)\int_0^1 K^2\Big(\frac{t-s}{b_n}\Big)E\Big[\frac{\alpha_r^{(n)}(s)}{(Y_r^{(n)}(s))^4}\Big]\,ds\Big\} \\
&\le \frac{n^2}{b_n^4 d_n^2}O(b_n d_n)O(b_n)O\Big(\frac{d_n}{n}\Big)^3 \\
&= O\Big(\frac{1}{n b_n^2}\Big) \to 0.
\end{aligned}$$

This completes the proof.

References

Aalen, O. O. (1977). Weak convergence of stochastic integrals related to counting processes. Z. Wahrsch. verw. Gebiete 38, 261-277. Correction: 1979, vol. 48, 347.

Aalen, O. O. (1978). Nonparametric inference for a family of counting processes. Ann. Statist. 6, 701-726.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. Ann. Statist. 10, 1100-1120.

Bass, R. F. (1988). Probability estimates for multiparameter Brownian processes. Ann. Probab. 16, 251-264.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Tech. Report, Dept. of Statistics, University of California, Berkeley.

Bickel, P. J. and Wichura, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. Ann. Math. Statist. 42, 1656-1670.

Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.

Billingsley, P. (1986). Probability and Measure. Second edition, Wiley, New York.

Cairoli, R. (1970). Une inégalité pour martingales à indices multiples et ses applications. Séminaire de probabilités IV, Lecture Notes in Math. 124, 1-27. Springer-Verlag, Berlin,

Cox, D. R. (1972). Regression models and life tables (with discussion). J. Roy. Statist. Soc. B 34, 187-220.

Dabrowska, D. M. (1987). Nonparametric regression with censored survival time data. To appear in Scand. J. Statist.

Dellacherie, C. and Meyer, P.-A. (1982). Probabilities and Potential B. North Holland, Amsterdam.

Hall, W. J. and Wellner, J. A. (1980). Confidence bands for a survival curve with censored data. Biometrika 67, 133-143.

Hastie, T. J. and Tibshirani, R. J. (1986). Generalized additive models (with discussion). Stat. Sci. 1, 297-310.

Ito, K.(1951). Multiple Wiener integral. J. Math. Soc. Japan 3, 157-169.

McKeague, I. W. (1988). A counting process approach to the regression analysis of grouped survival data. To appear in Stoch. Process. Appl.

McKeague, I. W. and Utikal, K. J. (1987). Inference for a nonlinear semimartingale regression model. Tech. Report, Dept. of Statistics, Florida State University, Tallahassee.

McKeague, I. W. and Utikal, K. J. (1988). In preparation.

Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. Ann. Math. Statist. 42, 1285-1295.

O'Sullivan, F. (1986a). Nonparametric estimation in the Cox proportional hazards model. Tech. Report, Dept. of Statistics, University of California, Berkeley.

O'Sullivan, F. (1986b). Relative risk estimation. Tech. Report, Dept. of Statistics, University of California, Berkeley.

Rebolledo, R. (1980). Central limit theorems for local martingales. Z. Wahrsch. verw. Gebiete 51, 269-286.

Thomas, D. (1983). Nonparametric estimation and tests of fit for dose response relations Biometrics, 39, 263-268.

Tibshirani, R. J. (1984). Local likelihood estimation. Tech. Report and unpublished Ph.D. dissertation, Dept. of Statistics, Stanford University.

Wong, E. and Zakai, M. (1974). Martingales and stochastic integrals for processes with a multidimensional parameter. Z. Wahrsch. verw. Gebiete 51, 109-122.

Ylvisaker, D. (1968). A note on the absence of tangencies in Gaussian sample paths. Ann. Math. Statist. 39, 261-262.